APPENDIX A: Document Type Definition of Multimedia Integration Description Scheme

MM_integration_ds.dtd

5

```
<!-- Multimedia Integration Description Scheme -->

<!ELEMENT MM_Stream ( MM_Object_Set, Object_Hierarchy*,
Entity_Relation_Graph* )>
<!ATTLIST MM_Stream
        id ID #IMPLIED>

<!-- This is how external DTDs are included in the current DTD -->

<!-- External Video DS DTD -->
<!ENTITY % Video_DS SYSTEM "video_ds.dtd">
%Video_DS;

<!—External Audio DS DTD  -->
<!ENTITY % Audio_DS SYSTEM "audio_ds.dtd">
%Audio_DS;

<!—External Text DS DTD  -->
<!ENTITY % Text_DS SYSTEM "text_ds.dtd">
%Text_DS;

<!—External Synthetic DS DTD -- >
<!ENTITY % Synthetic_DS SYSTEM "synthetic_ds.dtd">
%Synthetic_DS;

<!—External Image DS DTD  -->
<!ENTITY % Image_DS SYSTEM "image_ds.dtd">
%Image_DS;

<!ELEMENT MM_Object_Set ( MM_Object+ )>

<!ELEMENT MM_Object ( Media_Object_Set, Object_Hierarchy*,
Entity_Relation_Graph*,
                        MM_Obj_Media_Features?, MM_Obj_Semantic_Features?,
                        MM_Obj_Temporal_Features? )>

<!ATTLIST MM_Object
        Object_Type        (LOCAL|GLOBAL)        #REQUIRED
```

```
        id              ID                  #IMPLIED
        Object_Ref      IDREF               #IMPLIED
        Object_Node_Ref IDREFS              #IMPLIED
        Entity_Node_Ref IDREFS              #IMPLIED>

<!ELEMENT Media_Object_Set ( Audio_Object | Image_Object | Video_Object |
Text_Object
                        Synthetic_Object )+ >

<!-- The object hierarchy and the entity relation graph are defined in the Image DS
(Proposal #480). We include them in this DTD for convenience -->
<!-- Object hierarchy element -->
<!-- The attribute type is the hierarchy binding type -->
<!ELEMENT Object_Hierarchy ( Object_Node )>
<!ATTLIST Object_Hierarchy
        id          ID          IMPLIED
        type        CDATA       #IMPLIED>

<!ELEMENT Object_Node ( Object_Node* )>
<!ATTLIST object_node
        id          ID          #IMPLIED
        Object_Ref  IDREF       #REQUIRED>

<!-- Entity relation graph element-->
<!-- Possible types of entity relations and entity relation graphs:
        - Spatial: topological, directional
        - Temporal: topological, directional
        - Semantic -->
<!ELEMENT Entity_Relation_Graph ( Entity_Relation+ )>
<!ATTLIST Entity_Relation_Graph
        id          ID          #IMPLIED
        type        CDATA       #IMPLIED>

<!ELEMENT Entity_Relation ( Relation, (Entity_Node | Entity_Node_Set |
Entity_Relation)* )>
<!ATTLIST Entity_Relation
        type        CDATA       #IMPLIED>

<!ELEMENT Entity_Node (#PCDATA)>
<!ATTLIST Entity_Node
        id          ID          #IMPLIED
        Object_Ref  IDREF       #REQUIRED>

<!ELEMENT Entity_Node_Set ( Entity_Node+ )>
```

```
<!ELEMENT Relation Temporal_Parallel| Temporal_Sequential |
                Spatial_Alignment | Spatial_Arrangement |
                Semantic_Relation | code)*>
<!ATTLIST    Relation
             With_Respect_To    IDREF    #IMPLIED  >

<!ELEMENT    Temporal_Parallel      EMPTY >
<!ELEMENT    Temporal_Sequential  EMPTY >
<!ATTLIST    Temporal_Sequential
             Pattern   (EXACT|DELAY|PRIOR)     "EXACT" >

<!ELEMENT    Spatial_Alignment  EMPTY  >
<!ATTLIST    Spatial_Alignment
             Pattern   (Left_Align | Right_Align |
                    Top_Align | Bottom_Align)      "Left_Align"
             At_Time        CDATA              #IMPLIED  >

<!ELEMENT    Spatial_Arrangement ( Spatial_Relevance |  Spatial_Positioning )? >
<!ATTLIST    Spatial_Arrangement
             At_Time   CDATA    #IMPLIED  >
<!ELEMENT    Spatial_Relevance  EMPTY  >
<!ATTLIST    Spatial_Relevance
             Pattern  (Top_Of | Bottom_Of | Left_Of | Right_Of |
                    Upper_Left_Of | Upper_Right_Of |Lower_Left_Of |
                    Lower_Right_Of  |
                    Adjacent_To | Neighboring_To | Near_By |
                    Within | Contained_In )                  "Top_Of" >

<!ELEMENT    Spatial_Positioning  EMPTY  >
<!ATTLIST    Spatial_Positioning
             Horizontal_Shift       CDATA  #IMPLIED
             Vertical_Shift         CDATA  #IMPLIED  >

<!ELEMENT    Semantic_Relation  (Keywords* | code*)?  >


<!ELEMENT MM_Obj_Media_Features ( Data_Location?, Scalable_Representation?,
                                Modality_Transcoding? )>

<!ELEMENT MM_Obj_Semantic_Features ( Text_Annotation?, Keywords? )>

<!ELEMENT MM_Obj_Temporal_Features ( Duration? )>
```

```
<!ELEMENT Keywords ( Word*, Code* ) >
<!ATTLIST  Keywords
           No-Words          CDATA                        #REQUIRED
           Language          CDATA                        "English"
           Extraction_Manner  (AUTOMATIC|MANUAL)          "MANUAL" >

<!ELEMENT Duration    (
           Image_Duration?,
           Audio_Duration?,
           Video_Duration?,
           Text_Duration?,
           Synthetic_Duration? ) >
<!ATTLIST  Duration
           Synchronized_Overall_Duration CDATA #IMPLIED >
<!ELEMENT Image_Duration      ( Time ) >
<!ELEMENT Audio_Duration      ( Time ) >
<!ELEMENT Video_Duration      ( Time ) >
<!ELEMENT Synthetic_Duration  ( Time ) >
<!ELEMENT Text_Duration       ( Time? | Alignment? ) >
<!ELEMENT Alighment           EMPTY >
<!ATTLIST  Alignment
           with (IMAGE|AUDIO|VIDEO|SYNTHETIC)       "AUDIO" >

<!ELEMENT Data_Location(
           Image_Location?,
           Audio_Location?,
           Video_Location?,
           Text_Location?,
           Synthetic_Location?) >
<!ELEMENT Image_Location      (location) >
<!ELEMENT Audio_Location      (location) >
<!ELEMENT Video_Location      (location) >
<!ELEMENT Text_Location       (location) >
<!ELEMENT Synthetic_Location  (location) >


<!ELEMENT Scalable_Representation    (
           Static_Sampled?,
           Dynamic_Condensed?) >

<!ELEMENT  Static_Sampled (
           Image_Condensed?,
           Video_Static_Pictures?,
           Audio_Clips?,
```

```
                    Synthetic_Pictures?) >
      <!ELEMENT  Image_Condensed  ( Location*, Image_Scl*, Image_Subsampling* ) >
      <!ELEMENT  Image_Subsampling ( Image_Subsampling_Para, code)* >
      <!ELEMENT  Image_Subsampling_Para EMPTY >
5     <!ATTLIST  Image_Subsampling_Para
                    Scheme        CDATA  #REQUIRED
                    Spatial_Rate  CDATA  #REQUIRED
                    Frame_Size    CDATA  #IMPLIED >
      <!ELEMENT Video_Static_Pictures (Key_Frame* ) >
10    <!ATTLIST  No_KFs              CDATA #REQUIRED >
      <!ELEMENT  Audio_Clips (Audio_Object* | Audio_Hierarchy*) >
      <!ATTLIST  No_Clips            CDATA #REQUIRED >
      <!ELEMENT  Synthetic_Pictures (Key_Frame)* >
      <!ATTLIST  No-KFs              CDATA #REQUIRED >
15
      <!ELEMENT  Dynamic_Condensed (
                    Visual_Condensed?,
                    Audio_Condensed?,
                    Text_Condensed?,
20                  Synthetic_Condensed?) >

      <!ELEMENT  Visual_Condensed ( Location*, Video_Scl*, Video_Subsampling* ) >
      <!ELEMENT  Video_Subsampling     ( Video_Subsampling_Para, code)*  >
      <!ELEMENT  Video_Subsampling_Para EMPTY >
25    <!ATTLIST  Video_Subsampling_Para
                    Scheme         CDATA  #REQUIRED
                    Temporal_Rate  CDATA  #IMPLIED
                    Spatial_Rate   CDATA  #IMPLIED
                    Frame_Size     CDATA  #IMPLIED >
30    <!ELEMENT  Audio_Condensed ( Location*,
                                   Audio_Compressed*,
                                   Audio_Subsampling*,
                                   Audio_Timescaled* ) >
      <!ELEMENT  Audio_Compressed ( Audio_Compress_Para, code )*  >
35    <!ELEMENT  Audio_Compress_Para  EMPTY >
      <!ATTLIST  Audio_Compress_Para
                    Scheme  CDATA  #REQUIRED
                    Bitrate CDATA  #IMPLIED >
      <!ELEMENT  Audio_Subsampling ( Audio_Subsampling_Para, code)* >
40    <!ELEMENT  Audio_Subsampling_Para EMPTY >
      <!ATTLIST  Audio_Subsampling_Para
                    Scheme         CDATA  #REQUIRED
                    Temporal_Rate  CDATA  #IMPLIED >
      <!ELEMENT  Audio_Timescaled  (Audio_Timescale_Para, code)* >
```

```
<!ELEMENT    Audio_Timescale_Para  EMPTY >
<!ATTLIST    Audio_Timescale_Para
             Scale_Rate        CDATA  #REQUIRED >


<!ELEMENT    Text_Condensed  (Text_Abstract*) >
<!ELEMENT    Text_Abstract  (Location*, (Text_Abstract_Para, code)*)?  >
<!ATTLIST    Text_Abstract
             Length_In_Words      CDATA                  #IMPLIED
             Duration_In_Seconds  CDATA                  #IMPLIED
             Language             CDATA                  "English"
             Generation_Mode      (AUTOMATIC|MANUAL)     "MANUAL" >
<!ELEMENT    Text_Abstract_Para  EMPTY >
<!ATTLIST    Text_Abstract_Para
             Length_In_Words      CDATA                  #IMPLIED
             Language             CDATA                  "English" >


<!ELEMENT    Synthetic_Condensed  (Synthetic_Location*, (Synthetic_Condense_Para,
code)*) >
<!ELEMENT    Synthetic_Condense_Para  EMPTY >
<!ATTLIST    Synthetic_Condense_Para
             Spatial_Rate         CDATA                  #REQUIRED
             Temporal_Rate        CDATA                  #REQUIRED
             Frame_Size           CDATA                  #REQUIRED
             Bitrate              CDATA                  #IMPLIED >

<!-- Multimedia Integration DS End -->
```